CTPP 2000 Status Report

Census 2000 Place-of-Work Coding
By Clara Reschovsky, U.S. Census Bureau

The Census 2000 Place-of-Work (POW) coding operation started in early January 2001 and is scheduled to run through July of this year. At the Census Bureau's National Processing Center in Jeffersonville IN, clerks are hard at work trying to resolve the most difficult of the long-form POW responses from the Census 2000 questionnaires. It is a massive operation involving nearly 600 clerks and staff working across two shifts.

To date, all of the nearly 19 million POW person records have been processed through step one of the coding operation, automated (computer-based) coding. The automated coding software, developed in-house by Census Geography Division, matches workplace address responses against address ranges extracted from TIGER, to determine the street side (block) of each respondent's workplace. If this match is unsuccessful, the employer name response from the questionnaire is matched against our national employer file. Overall for Census 2000, 96.6% of the records with POW responses were coded to a city or town (the Census place level) in automated coding, while 54.6% of the records were successfully coded to a street side (block) in this first pass. Records that were not matched with confidence during automated coding were sent to step two of the coding operation, Computer-Assisted Clerical Coding (CACC).

The idea behind CACC is that clerks can be trained to exercise judgment and make

choices that would be very difficult to program into a computer algorithm. Using additional custom software developed by Geography Division, the coders can view both the questionnaire responses and the results of their searches of the reference files on their computer screens. CACC consists of two separate operations, depending on the level of geography being coded: census place-level coding and block-level coding.

Place-level CACC clerks successfully coded the vast majority of the 3.4% of records not assigned to a census place during automated coding, sending only about 70,000 cases to step three of the coding process, referral coding.

Block-level CACC is the largest task in POW coding in terms of the number of records to be processed. We use the employer information in conjunction with the work address information to try and

isolate the work location. The four main challenges are respondents providing inadequate or illegible information, TIGER inadequacies, incomplete employer files, and optical character reader errors. This operation is about 90% finished, and the clerks have successfully resolved about 52% of records not block-coded during the automated phase. The remaining 48% not coded in Block CACC are sent on to referral coding.

Referral coding is the last attempt to code and is also split between place-level and block-level work. Place-level referral clerks have additional resources available to them in their research such as maps and geographical dictionaries. Place-level referrals have been finished for the entire country.

Block-level referral coding is by far the most time-consuming part of the POW coding operation. Clerks attempt to group responses that appear to be the same employer or location with each other into what we call clusters. They then research the clusters using paper maps, zip code directories, internet searches, telephoning employers directly to find a usable address, and ArcView. We have modified Work-UP, the ArcView extension many of you used to review and update our employer file, to be used by the block referral clerks. As you can well imagine, a lot of creative detective work is necessary to complete referral coding to the block level. During April the other coding work will be finished and the entire coding staff in the National Processing Center will be devoted to POW block referrals.

We are scheduled to complete the POW coding operation at the end of the summer 2001. The next step is to send the geographic codes for the workers back to general decennial census processing to be merged with the basic person records.
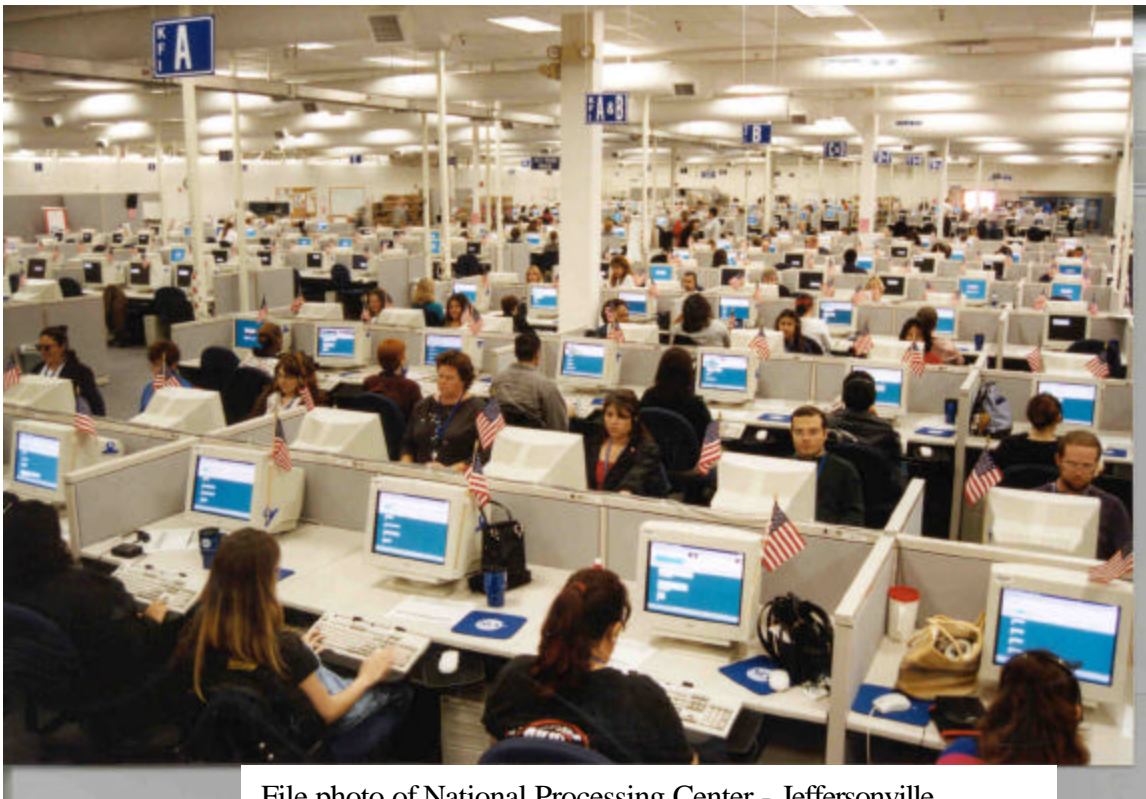
File photo of National Processing Center - Jeffersonville

First Small Geographic Data Released from Census 2000!

By Nanda Srinivasan, Cambridge Systematics Inc.

Small area (Census block) population counts by race and national origin are now available from the Census Bureau. These files are known as "Redistricting data" or the "PL-94-171 data". The files are based on 100 percent count and contain the most detailed information on the location of the total population by race and Hispanic/Latino origin, and population over the age of 18 (voting population) by race and Hispanic/Latino origin.

The PL-94-171 data constitute an important source for recent transportation planning concerns such as Environmental Justice (EJ) analysis because of the detailed race and ethnicity variables on the file. The file contains extensive data on race and ethnicity, so it can be used to analyze the concentrations of minority population groups.

The U.S. Census Bureau provides far more data at larger geographic units (such as tracts and block groups) in Sample File 3 (SF3), which will not be available until summer 2002.

For those areas that defined TAZs for TIGER 2000, a TAZ field is included in the file, so that users can aggregate blocks into TAZ summaries. You can easily overlay other TAZ level information (developed locally) over the PL-94-171 data.

To order the data for your region, please visit the Census Bureau website at: http://www.census.gov/cgi-bin/pubinfo/cdrequest. You can also get a copy of the CD-ROM from your Census State Data Center. For a list of State Data Centers, please visit: http://www.census.gov/pub/sdc/www/sdctxt.html.

For more information, please visit the Census Bureau website at: http://www.census.gov/clo/www/redistricting.html.

A method to transfer the PL-94-171 data to a GIS is posted on the CTPP website at http://www.mcs.com/~berwyned/census/products/pl2gis.html.

Disclaimer:
The CTPP Working Group is primarily involved in the creation, dissemination, and optimum use of the Census Transportation Planning Package. But, we also want to do our best in making the transportation community aware of other decennial data products. We, however, do not have the staff resources to provide technical assistance to MPO or State DOT staff on other products.

Please contact your Census Bureau State Data Center for assistance or questions on other Census products.

Meeting of the TRB Subcommittee on Census Data for Transportation Planning
by Ed Limoges, Secretary

The Subcommittee on Census Data for Transportation Planning A1D08(1) met on Monday, January 8, 2001 during the Transportation Research Board (TRB) Annual Meeting in Washington DC. The topics covered included:

Chair's Report
Chair Ed Christopher, Bureau of Transportation Statistics, declared that this has been a busy year for the Subcommittee, as well as for the CTPP Planning Group (17 meetings, four issues of the *CTPP Status Report)*. Both TAZ-UP and Work-UP were completed by the participating MPOs and states. The CTPP now has a brochure. Outreach efforts have been expanded; during 2000, various members have traveled nationwide to give sessions or presentations at more than ten professional meetings. The Public Use Microdata Sample (PUMS) issues of sample size and confidentiality compliance emerged and were resolved.

The subcommittee's website has grown and has become a stable source of information on the CTPP. It is being moved to a FHWA server. We will keep the list serve — currently about 650 subscribers. We are always looking for examples of CTPP data users, please pass along names and contact information. Other ongoing efforts include the archiving of 1980 UTPP files, and the compilation of a CTPP history.

Census Coding in Jeffersonville
Phil Salopek, Census Bureau, reported on the progress of automated workplace coding, noting that the biggest constraint is the quality of questionnaire responses.

Urbanized Area Definitions
Joe Marinucci, Census Bureau, reported on the largely automated process that will be used to demarcate urbanized areas (population 50,000 and over) and urban clusters (population 2,500 to 49,999). Criteria including rules on minimum densities, linking noncontiguous areas, and closing indentations has been published in the *Federal Register* on March 28, 2001 (http://www.mcs.com/~berwyned/census/notes/ua_notice032801.pdf) Delineation of urbanized areas is planned to be completed by January 2002.

American Fact Finder
Barbara Aldrich, Census Bureau, gave a presentation on using the American FactFinder (AFF). The AFF provides Internet access to Census Bureau data from the 1990 Decennial Census, Census 2000, the 1997 Economic Census, and the American Community Survey. It can be accessed at http://factfinder.census.gov.

CTPP Table Summary
Nandu Srinivasan, Cambridge Systematics Inc., presented a handout on and discussed the proposed CTPP standard tabulations.

American Community Survey (ACS)
Larry Mcginn, Census Bureau, reported on the ACS, an ongoing survey using a rolling sample, that replaces the census long form. Several working groups are considering the question content of the ACS, with a federal agencies group leaning toward increasing questions, and a congressional group leaning the opposite way. The Census 2000 long form question about time of departure from home generated some negative comments.

Attendees felt it was critical to keep reminding those involved in discussions of a future ACS that transportation planning must have data at the small area level.

Related issues included changes in TAZ boundaries over time, geocoding ACS journey to work questions, periodic updating of the reference files, and funding issues for a CTPP based on ACS.

TRB Subcommittee: Message from the Chair:
American Community Survey (ACS) – An Update

Because of Congressional concerns about privacy and intrusiveness, the Office of Management and Budget (OMB) has convened an Interagency Committee on the ACS. Right now, the Census Bureau (CB) is conducting a test of the ACS (1999-2001) from which the results will be compared to decennial census long form data. The full, nationwide, implementation of ACS is currently scheduled to begin in 2003.

As a first step, OMB asked each federal agency to select those questions/items in the current ACS form that they feel responsible for and identify the legislative basis for inclusion of each question. That is, instead of assuming that all the questions on the 2000 decennial Census Long Form would be included in the full ACS, we would start over from "square one." USDOT submitted a package to OMB detailing the legislative requirements for the transportation-related questions. A PDF version of this document can be found at http://www.fhwa.dot.gov/ctpp/articles/omb2001.html. A hard copy can be obtained by emailing ctpp@fhwa.dot.gov.

As noted in the USDOT package, there are no federal laws that "mandate" the CB to collect and provide the data used by the transportation community (other than population). However, the CB data has evolved into the best, and in many cases, the only source of information for meeting the myriad of federal planning and Congressional requirements.

Specific to the transportation items there is one that is somewhat problematic--Departure time from home to go to work.

There have allegedly been citizen complaints about this question, and the ACS staff (Larry McGinn) is feeling pressure to drop it from the questionnaire. Chuck Purvis, chair of the TRB Urban Data Committee (A1D08) surveyed the transportation community to identify the importance of the departure time question and is preparing his comments. In general there is strong support for this question and it is in fact being used. Its usage varies from building and analyzing peak-period travel models to congestion analysis to rideshare marketing. The bottom line is that this question will become even more important in the future as we are called on to more effectively manage our transportation systems.

Stay tuned for more ACS developments…

Ed Christopher,
Chair, TRB Sub-committee on Census Data for Transportation Planning

---

TRB Subcommittee Meeting on Census Data (Continued from Page 4)

CTPP Outreach and Software:
Elaine Murakami, Federal Highway Administration, reported that work is proceeding to create software for special tabulations. Typically, CTPP users want specific characteristics for all TAZs. Work will soon begin to produce national tabulations of the *Commuting in America* type. The contract for a video, DVD or internet based course and a brochure, should be signed shortly.

Applications of the 1990 Census Public-Use Microdata Sample (PUMS) Files

By Nanda Srinivasan, Cambridge Systematics Inc.

Public Use Microdata Sample (PUMS) files are a small sample of individual records from the census long form. To prevent individual disclosure, the geographic unit of reporting used in PUMS is large.

PUMS contains records from 5 percent of all housing units. Compared to other census products such as the redistricting file (based on 100% count released at the block level), or the Sample File 3 (based on 16% sample released at the tract/block group level), PUMS contains lesser geographic specificity, and sample size. Figure 1 depicts a comparison of PUMS with other census data, and travel survey data.
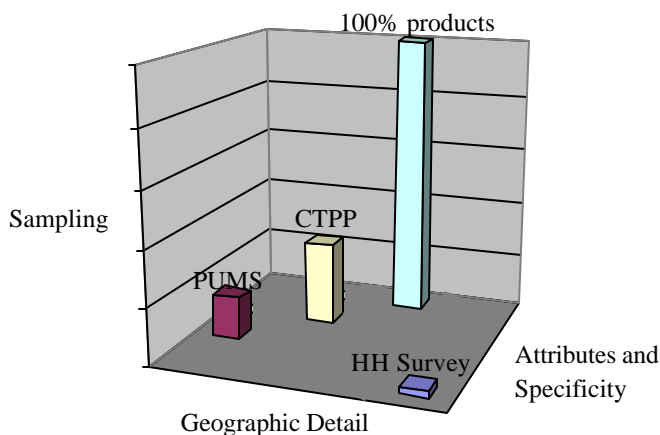

Figure 1: Dataset Comparison

Testing Model Specifications and Transferability

Jim Ryan and Gregory Han (Parsons Brinckerhoff) used PUMS to develop a vehicle-ownership model for Honolulu, Hawaii. The model was developed for Honolulu, and then applied for other cities such as Kansas City, Atlanta, San Francisco,

and New York to test for behavioral content. Thus, PUMS can be used as data source to check for transferability of models as it is consistent in content and format across metropolitan areas. The study is documented in the article, "Vehicle-ownership model using family structure and accessibility application to Honolulu, Hawaii" in Transportation Research Record 1676.

Another example for a vehicle ownership model is the nested workers in household / auto ownership choice model (WHHAO) developed by Metropolitan Transportation Commission (MTC), the MPO for the San Francisco area. This model splits the households residing in a travel analysis zones into households by three workers in household levels (0, 1, 2+ workers/HH) by three auto ownership levels (0, 1, 2+ vehicles/HH). The input market segmentation to the WHHAO model is households by household income quartile. This means that the outputs of the WHHAO model application are the number of households in each travel zone stratified by household income (4) by workers in household (3) by auto ownership level (3), or four market segmentations *into* the model, thirty-six market segmentations coming *out* of the model application.

One of the input variables to the WHHAO model is average household size. Because low income households (less than $25,000 per year) are smaller on average than higher income households, average household size for the transportation analysis zone need to be adjusted for different income quartiles. PUMS data is used to estimate average household size stratified by four household income categories. These county-level adjustment factors are then applied to smaller zones.

Table 1 shows average household size adjusted by income for the North Beach Transportation Analysis Zone. The average

household size in the TAZ = 1.719 persons per household (obtained from CTPP). This is adjusted by county derived factor for different income quartiles.

Table 1: Average Household Size by Income

	County Adjustment Factor	Average Household size adjusted by income
Low	0.792	1.361
Medium-Low	0.994	1.709
Medium-High	1.175	2.020
High	1.296	2.228

Thus, the subgroup mean household size in the North Beach ranges from 1.361 to 2.228 persons per household instead of just 1.719 obtained from CTPP.

These adjustment factors used in the market segmentation process could also be developed using data from local household travel surveys. Because these travel surveys are small, the reliability of these factors is reduced, compared to decennial census data. For example, the 1990 Bay Area household travel survey provides sample data on 10,800 households. The 1990 Census PUMS 5-percent sample includes disaggregate data on 108,500 Bay Area households. For more information on how MTC uses Census PUMS, please visit www.mtc.ca.gov.

Sample Expansion

PUMS can be used in survey sample expansion to control for differential response rates by socio-economic characteristics and by large geographic areas. For example, Central Planning Transportation Staff (CTPS), the MPO for the Boston area used PUMS for expanding their 1990 household survey data. PUMS was used along with CTPP data to generate some 3- and 4-way cross tabulations that were not available in the CTPP. For

example, the CTPS wanted to use a three-way table containing household size, number of workers in household, and household income to expand their survey data. However, the 1990 CTPP did not contain this table, but it contained two way tables with household income and household size; and number of workers in household and household income.

Note: For CTPP 2000, a three way table containing household size, income and number of workers in household has been specified.

Travel Demand Modeling

1990 Census PUMS is being used in micro-simulation models. For example, TRANSIMS is a micro-simulation model being developed by US DOT and EPA. It depends on Census PUMS at the front end to create "synthetic" household information. The "Population Synthesizer" routine in TRANSIMS takes in various types of census data to generate synthetic households, individuals, and vehicles through a series of 6 steps. To evaluate how TRANSIMS uses PUMS please visit http://transims.tsasa.lanl.gov/ MSWord_Files/U-Vol3-Chap2-PopSynth-Mar1.doc.

Another example of a micro-simulation is the Short-Range Transportation Evaluation Program (STEP) model. STEP is run on actual or forecast data on household socio-economic characteristics, population and employment, and transportation system characteristics. Data are taken from a regional household survey, and alternatively, from Census PUMS. STEP was developed by Greig Harvey and Elizabeth Deakin (University of California, Berkeley) and is published as Appendix B in "Guidance on the Use of Market Mechanisms to Reduce Transportation Emissions", US E.P.A., Washington, DC, Fall 1996.

CTPP Hotline – 202-366-5000

FHWA:
Elaine Murakami
PH: 202-366-6971
 (206-220-4460 in Seattle)
FAX: 202-366-7660
Email: elaine.murakami@fhwa.dot.gov

Nanda Srinivasan
PH: 202-366-5021
FAX: 202-366-7742
Email: nanda.srinivasan@fhwa.dot.gov

BTS:
Ed Christopher (Census Subcommittee
Chair)
PH: 202-366-0412
FAX: 202 366-3640
Email: berwyned@mcs.com

FTA:
Jim Ryan
PH: 202-366-0954
FAX: 202-493-2478
Email: james.ryan@fta.dot.gov

Sarah Clements
PH: 202-366-4967
FAX: 202-493-2478
Email: sarah.clements@fta.dot.gov

AASHTO:
Deborah Buchacz
PH: 202-624-5839
FAX: 202-624-5806
Email: debbieb@aashto.org

Census Population Division (JTW
questions) (Please note change in addresses)
Phil Salopek
PH: 301-457-2454
Fax: 301-457-2481
Email: phillip.a.salopek@census.gov

Clara Reschovsky
PH: 301-457-2454
FAX: 301-457-2481
Email: clara.a.reschovsky@census.gov

TRB Committees
Ed Christopher (Census Subcommittee) -
See under BTS

Ron Tweedie (State Data Committee Chair)
PH: 518-439-4463
Email: rtweedie@NYCAP.rr.com

Chuck Purvis (Urban Data Committee
Chair)
PH: 510-464-7731
FAX: 510-464-7848
Email: cpurvis@mtc.ca.gov

Some Important Websites:

NEW **CTPP Website:** The new, easy to remember CTPP Website address is:
http://www.dot.gov/ctpp
TRB Census Data Subcommittee:
http://www.mcs.com/~berwyned/census/
American Community Survey (ACS):
http://www.census.gov/acs/www/